

SOURCE AND ACCURACY OF THE DATA

Source of Data

The data for this report, which cover a wide range of topics and years, came from the Current Population Survey (CPS), the Survey of Income and Program Participation (SIPP), the American Housing Survey (AHS), the Census Bureau's Estimates Program, and the decennial censuses conducted by the Census Bureau. The surveys' estimation procedure adjusts weighted sample results to agree with independent estimates of the civilian noninstitutional population of the United States by age, sex, race, Hispanic/non-Hispanic ancestry, and state of residence.

The independent estimates are calculated based on information from four primary sources: the 1990 Decennial Census of Population and Housing, statistics on births, deaths, immigration, and emigration; statistics on the size of the Armed Forces; and starting in 1994, an adjustment for undercoverage in the 1990 census. The estimation procedure for 1994 and later years used independent estimates based on the most recent decennial census at that time. (Data in some sections are revised for years prior to 1994.) This change in independent estimates had relatively little impact on summary measures, such as medians and percent distributions, but did have a significant impact on levels. For example, use of the 1990-based population controls resulted in about a 1-percent increase in the civilian noninstitutional population and in the number of families and households. Thus, estimates of levels for 1994 and later years will differ from those for earlier years by more than what could be attributed to actual changes in the population. These differences could be disproportionately greater for certain population subgroups than for the total population.

The estimation procedures for CPS, SIPP, and AHS data are discussed in more detail in the publications cited in Appendix A of this report.

Reliability of Estimates

Since the CPS, SIPP, and AHS estimates come from samples, they may differ from the figures from a complete census using the same questionnaires, instructions, and enumerators. This possible variation in the estimates due to sampling error is known as "sampling variability." A sample survey estimate has two types of error: sampling and nonsampling. The accuracy of an estimate depends on both types of error. The nature of the sampling error is known given the survey design. The full extent of nonsampling error, however, is unknown.

To estimate the standard error of a CPS estimate, the Census Bureau uses replicated variance estimation methods. These methods primarily measure the magnitude of sampling error. However, they do measure some effects of nonsampling error as well. They do not measure systematic biases in the data due to nonsampling error. (Bias is the average of the differences, over all possible samples, between the sample estimates and the desired value.)

Since the full extent of nonsampling error is unknown, one should be particularly careful when interpreting results based on small differences between the estimates. Even a small amount of nonsampling error can cause a borderline difference to appear significant or not, thus distorting a seemingly valid hypothesis test. Caution should also be used when interpreting results based on a relatively small number of cases. Summary measures probably do not reveal useful information when computed on a base smaller than 75,000.

Sampling Error

Standard errors are not given in this report because of the wide range of topics included and the wide variety of data sources. Standard errors may be found in the publications that are noted at the end of most sections and in Appendix A or by contacting the subject specialist provided at the end of each section.

Nonsampling Variability

As in any survey work, the results are subject to errors of response and nonreporting in addition to sampling variability. Nonsampling errors can be attributed to many sources; for example, inability to obtain information about all cases, definitional difficulties, differences in the respondent interpretation of questions, respondent inability or unwillingness to provide correct information, respondent inability to recall information, errors made in collection such as recording or coding data, errors made in processing the data, errors made in estimating values for missing data, and failure to represent all units with the sample (undercoverage).

Comparability of Data

Data obtained from sample surveys and other sources are not entirely comparable. This results from differences in interviewer training and experience and in differing survey processes. This is an example of nonsampling variability not reflected in the standard errors. Therefore, caution should be used in comparing results from different sources.

A number of changes were made in data collection and estimation procedures beginning with the January 1994 CPS. The major change was the use of a new questionnaire. The questionnaire was redesigned to measure the official labor force concepts more precisely, to expand the amount of data available, to implement several definitional changes, and to adapt to a computer-assisted interviewing environment. The March supplemental income questions were also modified for adaptation to computer-assisted interviewing, although there were no changes in definitions and concepts. Due to these and other changes, caution should be used when comparing estimates from data collected before 1994 with estimates from data collected in 1994 or later. For a description of these changes and the effect they had on the data, see the publications noted in Appendix A and at the end of most sections.

The April 1, 1990, census population was about 1.5 million less than the estimate for the same date obtained by carrying forward the 1980 census population on births, deaths, legal international migration, and the net migration of U.S. citizens across national boundaries. There are several possible explanations for the difference, or "error of closure," including a larger net underenumeration in the 1990 census, and duplications and erroneous enumeration in the 1980

census. For a detailed discussion, see J. Gregory Robinson, Bashir Ahmed, Prithwis Das Gupta, and Karen A. Woodrow, "Estimating Coverage in the 1990 United States Census," *Journal of the American Statistical Association*, 88, No. 423 (1993):1061-1071.

This report includes data for three different population universes: resident population (census universe): civilian noninstitutional population, plus Armed Forces living off post or with their families on post (SIPP and March CPS universe), as well as the universe of housing units. The estimated civilian noninstitutional population on July 1, 1999 was 267,703,000 (Table B-1). This population is not adjusted for estimated net underenumeration in the 1990 census. However, it incorporates a small increase (8,429 persons) in the census-base population from count resolution correction processed through 1999. Housing unit undercoverage was about 1.9 percent for the 1999 AHS.

While civilian noninstitutional population has been adopted as the universe for many sample surveys, the data in Tables B-1 and B-2 are not consistent with results of current surveys conducted by the Census Bureau through the end of 1993, including the CPS which were calibrated to 1980 or earlier census-based projections. Current estimates for dates from January 1, 1994 onward are not consistent with the results of those surveys, including the CPS, which are calibrated to projections that have been adjusted for estimated net underenumeration based on the 1990 Post Enumeration Survey.

The resident Armed Forces and the institutional population differ greatly from the resident population in age-sex structure (Table B-2) on July 1, 1999, males 18 to 64 years old constituted 85.8 percent of the resident Armed Forces population, compared with 30.4 percent of the institutional population, and females 65 years and over constituted 34.3 percent of the institutional population compared with 7.4 percent of the resident population. However, these two groups together (resident Armed Forces and institutional population) accounted for only about 1.8 percent of the resident population. As a result, the civilian noninstitutional population (which accounts for 98.2 percent of the resident population) has an age-sex structure very similar to that of the resident population. Similarly, the social and economic characteristics of the resident Armed Forces and the institutional population could differ greatly from those of the resident population, despite relatively small differences between the characteristics of the resident population and of the civilian noninstitutional population.

Table B-1.

Components of Selected Population Universes: July 1, 1999

(Numbers in thousands. These estimates are consistent with the 1990 census, as enumerated)

Population universe	Number	Percent
Resident population ¹	272,691	100.0
Resident Armed Forces.....	1,199	0.4
Civilian population ²	271,491	99.6
Institutional population ³	3,789	1.4
Noninstitutional population ³	267,703	98.2

¹Estimates of the U.S. resident population include people resident in the 50 states and the District of Columbia, but not in Puerto Rico. These estimates exclude the U.S. Armed Forces overseas, as well as civilian U.S. citizens whose usual place of residence is outside the United States.

²Civilian population estimates include U.S. residents not in the active duty Armed Forces. The difference between resident population plus Armed Forces overseas and civilian population is the worldwide Armed Forces population.

³The institutional population is estimated from proportions of the total residing in institutions at the time of the 1990 census, applied to current estimates of the total population by age and sex. The civilian noninstitutional population is computed as the difference between the civilian population and the institutional population.

Source: U.S. Census Bureau, 1999 estimates.

Table B-2.

Selected Population Universes by Sex and Broad Age Groups: July 1, 1999

(These estimates are consistent with the 1990 census as enumerated)

Population universe and age	Population			Percent of population universe		
	total	Male	Female	Total	Male	Female
RESIDENT POPULATION ¹						
Total.....	272,690,813	133,276,559	139,414,254	100.0	48.9	51.1
Under 18.....	70,199,435	35,960,621	34,238,814	25.7	13.2	12.6
18 to 64.....	167,951,353	83,005,866	84,945,487	61.6	30.4	31.2
65 and older.....	34,540,025	14,310,072	20,229,953	12.7	5.2	7.4
RESIDENT ARMED FORCES						
Total.....	1,199,338	1,031,038	168,300	100.0	86.0	14.0
Under 18.....	2,853	2,221	632	0.2	0.2	0.1
18 to 64.....	1,196,485	1,028,817	167,668	99.8	85.8	14.0
65 and older.....	—	—	—	—	—	—
INSTITUTIONAL POPULATION						
Total.....	3,788,725	2,170,084	1,618,641	100.0	57.3	42.7
Under 18.....	180,154	135,380	44,774	4.8	3.6	1.2
18 to 64.....	1,830,844	1,556,809	274,035	48.3	41.1	7.2
65 and older.....	1,777,727	477,895	1,299,832	46.9	12.6	34.3
CIVILIAN NONINSTITUTIONAL POPULATION ²						
Total.....	267,702,750	130,075,437	37,627,313	100.0	48.6	51.4
Under 18.....	70,016,428	35,823,020	34,193,408	26.2	13.4	12.8
18 to 64.....	164,924,024	80,420,240	84,503,784	61.6	30.0	31.6
65 and older.....	32,762,298	13,832,177	18,930,121	12.2	5.2	7.1

— Not applicable.

¹Estimates of the U.S. resident population include people resident in the 50 states and the District of Columbia, but not Puerto Rico. These estimates exclude the U.S. Armed forces overseas, as well as civilian U.S. citizens whose usual place of residence is outside the United States.

²The institutional population is estimated from proportions of the total population residing in institutions at the time of the 1990 census, applied to current estimates of the total population by age and sex. The civilian noninstitutional population is computed as the difference between the civilian population and the institutional population.

Source: U.S. Census Bureau, 1999 estimates.